Université Lyon 1 Mathématiques

L3 Probabilités et Statistiques

Examen du 9 mai 2017 – Corrigé

Exercice 1

1. On a pour $k \in \mathbf{N}$, $\mathbf{P}(N > k) = \mathbf{P}(X_0 = \cdots = X_k = Y_0 = \cdots = Y_k = 0) = (1/2)^{2(k+1)}$ par indépendance, formule valable aussi pour $k = -1$. On a donc
$$\mathbf{P}(N = k) = \mathbf{P}(N > k-1) - \mathbf{P}(N > k) = (1/4)^k - (1/4)^{k+1} = (3/4)(1/4)^k$$
ce qui montre que N suit une loi gémétrique de paramètre $3/4$.
2. On a $\{N = n, X_N = 1, Y_N = 0\} = \{X_0 = \cdots = X_{n-1} = 0, X_n = 1, Y_0 = \cdots = Y_n = 0\}$ donc l'événement a probabilité $2^{-2(n+1)}$.
3. On a donc $\mathbf{P}(X_N = 1, Y_N = 0) = \sum_{n=0}^{\infty} \mathbf{P}(N = n, X_N = 1, Y_N = 0) = \sum_{n=0}^{\infty} (\frac{1}{4})^{n+1} = 1/3$.
4. On a $\mathbf{P}(X_N = 0, Y_N = 0) = 0$ et $\mathbf{P}(X_N = 0, Y_N = 1) = \mathbf{P}(X_N = 1, Y_N = 1) = 1/3$ en raisonnant comme prédécemment.
5. On lance deux fois la pièce. Si le résultat pile/pile, on choisit 1 ; si c'est pile/face, on choisit 2 ; si c'est face/pile, on choisit 3 ; si c'est face/face, on recommence avec deux nouveaux lancers. On obtient ainsi les événements de probabibilté 1/3 considérés à la question précédente.

Exercice 2

1. On a par indépendance $\mathbf{E}|Y_n| = \mathbf{E}|X_n X_{n+1}| = \mathbf{E}|X_n| \times \mathbf{E}|X_{n+1}| < +\infty$.
2. Les variables aléatoires $(Y_{2n+1})_{n \geqslant 0}$ sont indépendantes (par groupement par paquets) de même loi et intégrables. Le résultat demandé découle de la loi forte des grands nombres. De même, les v.a. $(Y_{2n+1})_{n \geqslant 1}$ sont i.i.d.
3. Soit $\varepsilon > 0$. On a
$$\sum_{n=1}^{\infty} \mathbf{P}(|Y_n|/n \geqslant \varepsilon) = \sum_{n=1}^{\infty} \mathbf{P}(|Y_1/\varepsilon| \geqslant n) \leqslant \int_0^{\infty} \mathbf{P}(|Y_1/\varepsilon| \geqslant t)\,\mathrm{d}t = \mathbf{E}|Y_1/\varepsilon| < \infty$$
donc la série $\sum_{n=1}^{\infty} \mathbf{P}(|Y_n|/n \geqslant \varepsilon)$ converge. Par le lemme de Borel–Cantelli, on a p.s. $|Y_n|/n < \varepsilon$ pour n assez grand. C'est vrai simultanément p.s. pour tout $\varepsilon > 0$ rationnel, ce qui implique que (Y_n/n) converge p.s. vers 0.
4. Notons $Z_n = (Y_1 + \cdots + Y_n)/n$. La sous-suite (Z_{2n}) converge p.s. vers $\mathbf{E}Y_1$ comme conséquence de la question 2 (ajouter les 2 équations). Puisque $Z_{2n+1} = \frac{2n}{2n+1} Z_{2n} + \frac{Y_{2n+1}}{2n+1}$, la question 3 implique alors que la sous-suite (Z_{2n+1}) converge p.s. vers $\mathbf{E}Y_1$. On donc donc p.s. convergence à la fois de (Z_{2n}) et (Z_{2n+1}), donc de (Z_n), vers $\mathbf{E}Y_1$.

Exercice 3

1. Soit $0 < \varepsilon < 1$. Tout d'abord pour tout n, il existe K_n tel que $\mathbf{P}(|X_n| \geqslant K_n) \leqslant \varepsilon$ puisque $\bigcap_{K \in \mathbf{N}} \{|X_n| \geqslant K\} = \emptyset$. Notons X la limite de X_n, et soit x, y des points de continuité de F_X vérifiant $F_X(x) < \varepsilon/2$ et $F_X(y) > 1 - \varepsilon/2$. La caractérisation de la converge en loi par les fonctions de répartition implique que $\lim F_{X_n}(x) = F_X(x)$ et $\lim F_{X_n}(y) = F_X(y)$. Ainsi pour n assez grand (disons $n > n_0$), on a $F_{X_n}(x) < \varepsilon/2$ et $F_{X_n} > 1 - \varepsilon/2$, donc
$$\mathbf{P}(X_n \notin [x, y]) = \mathbf{P}(X_n < x) + \mathbf{P}(X_n > y) \leqslant \mathbf{P}(X_n \leqslant x) + 1 - \mathbf{P}(X_n \leqslant x) \leqslant \varepsilon.$$
On peut finalement choisir $K = \max(-x, y, K_1, \ldots, K_{n_0})$.
2. Soit $\varepsilon > 0$, et K donné par la question précédente. On a
$$\mathbf{P}(|X_n Y_n| \geqslant \varepsilon) \leqslant \mathbf{P}(|X_n| \geqslant K \text{ ou } |Y_n| \geqslant \varepsilon/K) \leqslant \mathbf{P}(|X_n| \geqslant K) + \mathbf{P}(|Y_n| \geqslant \varepsilon/K) \leqslant \varepsilon + \mathbf{P}(|Y_n| \geqslant \varepsilon/K)$$
et cette quantité est majorée par 2ε pour n assez grand puisque (Y_n) tend en probabilité vers 0.

Exercice 4

1. Le plus simple est de vérifier l'égalité des fonctions caractéristiques : pour tout $t \in \mathbf{R}$
$$\Phi_{(X+Y)/\sqrt{2}}(t) = \Phi_{X+Y}(t/\sqrt{2}) = \Phi_X(t/\sqrt{2})\Phi_Y(t/\sqrt{2}) = (\exp(-\sigma^2 t^2/4))^2 = \Phi_X(t)$$
2. Découle du fait que X et $\frac{X+Y}{\sqrt{2}}$ ont même espérance, et de la linéarité de l'espérance.
3. Par récurrence sur n.
4. Soit (X_n) des copies indépendantes de X, et $Y_n = (X_1 + \cdots + X_n)/\sqrt{n}$. Par le théorème central limite, la suite (Y_n) converge en loi vers une v.a. de loi $N(0, \sigma^2)$. La sous-suite (Y_{2^n}) converge aussi en loi vers la même limite ; comme cette sous-suite est constante en loi de même loi que X, il suit que $X \sim N(0, \sigma^2)$.

Exercice 5

1. $\int_0^\theta f_\theta = \frac{1}{\sqrt{\theta}} \int_0^\theta \frac{1}{2\sqrt{x}} \,\mathrm{d}x = \frac{\sqrt{\theta}}{\sqrt{\theta}} = 1$
2. (a) Par linéarite de l'espérance, $\mathbf{E}\tilde{\theta}_n = 3\mathbf{E}X_1 = 3\int_0^\theta x f_\theta(x)\,\mathrm{d}x = \frac{3}{2\sqrt{\theta}} \int_0^\theta \sqrt{x}\,\mathrm{d}x = \theta$. De plus, par indépendance de (X_n),
$$\mathbf{Var}\tilde{\theta}_n = \frac{9}{n^2}\left[\mathbf{Var}(X_1) + \cdots + \mathbf{Var}(X_n)\right] = \frac{9}{n}\mathbf{Var}X_1.$$
On calcule
$$\mathbf{E}X_1^2 = \int_0^\theta x^2 f_\theta(x)\,\mathrm{d}x = \frac{1}{2\sqrt{\theta}} \int_0^\theta x^{3/2}\,\mathrm{d}x = \theta^2/5.$$
On a donc $\mathbf{Var}X_1 = \mathbf{E}(X_1^2) - (\mathbf{E}X_1)^2 = \theta^2/5 - \theta^2/9 = 4\theta^2/45$. Finalement $\mathbf{Var}\tilde{\theta}_n = \frac{4\theta^2}{5n}$.

   (b) On applique le théorème central limite à la suite de v.a. (X_n), i.i.d. L^2 de moyenne $\theta/3$ et de variance $4\theta^2/45$. On a donc la convergence en loi de $(X_1 + \cdots + X_n - n\theta/3)/\sqrt{n}$ vers une v.a. de loi $N(0, 4\theta^2/45)$. De manière équivalente, $\sqrt{n}(\tilde{\theta}_n - \theta)$ converge en loi vers une v.a. de loi $N(0, 4\theta^2/5)$.
3. (a) Pour $t \in (0, \theta)$ $\mathbf{P}(\hat{\theta}_n \leqslant t) = \mathbf{P}(\forall i \in \{1, \cdots, n\}, X_i \leqslant t) = \mathbf{P}(X_1 \leqslant t)^n = (\sqrt{t/\theta})^n$.

   (b) On a p.s. $X_i < \theta$ pour tout i, donc $\hat{\theta}_n < \theta$. On a donc $\mathbf{E}\hat{\theta}_n < \theta$: $\hat{\theta}$ est biaisé.

   (c) Pour tout $\varepsilon > 0$, on a $\mathbf{P}(|\hat{\theta}_n - \theta| \geqslant \varepsilon) = \mathbf{P}(\hat{\theta}_n \leqslant \theta - \varepsilon) = (1 - \varepsilon/\theta)^{n/2} \to 0$ donc $\hat{\theta}$ est consistant.

   (d) Montrons que $\hat{\theta}$ a vitesse n : effet pour $x > 0$
$$\mathbf{P}(n(\hat{\theta}_n - \theta) < -x) = \mathbf{P}(\hat{\theta}_n < \theta - \frac{x}{n}) = \left(1 - \frac{x}{n\theta}\right)^{n/2}$$
qui tend vers $\exp(-x/2\theta)$ quand $n \to \infty$. On a donc
$$\lim_{n\to\infty} \mathbf{P}(n(\hat{\theta}_n - \theta) < x) = \begin{cases} \exp(x/2\theta) & \text{si } x < 0 \\ 1 & \text{sinon.} \end{cases}$$
Ainsi $n(\hat{\theta}_n - \theta)$ converge en loi vers une v.a. non constante, d'où le résultat.

   (e) L'estimateur $\hat{\theta}$ est préférable même s'il est biaisé car il a une vitesse supérieure (n au lieu de $\sqrt{n}$).